融合迁移学习和集成学习的自然背景下荒漠植物识别方法

王亚鹏[1,2]，曹姗姗[2,3]，李全胜[1]，孙　伟[2,3*]

（1. 新疆农业大学 计算机与信息工程学院，新疆乌鲁木齐 830052；2. 中国农业科学院农业信息研究所，北京 100081；3. 国家农业科学数据中心，北京 100081）

摘　要：［**目的/意义**］荒漠植物的准确识别是其认识和保护过程中不可或缺的任务，是荒漠生态研究与保护的基础。自然条件下野外荒漠植物图像的机器视觉自动分类识别可有效提升植物资源调查效率、降低人为主观因素影响，对荒漠植物的精准分类、多样性保护和资源化利用具有重要意义。［**方法**］以自然环境下的整株荒漠植物图像为研究对象，构建新疆干旱区荒漠植物图像数据集，以EfficientNet B0—B4网络为基础网络，提出一种融合迁移学习和集成学习的荒漠植物图像识别算法，并在公开数据集Oxford Flowers102上进行对比验证。［**结果和讨论**］基于EfficientNet B0网络的单一子模型的Top-1准确率最高可达93.35%，最低为92.26%，软投票Ensemble-Soft模型、硬投票Ensemble-Hard模型以及加权投票法集成的Ensemble-Weight模型的准确率分别为93.63%、93.55%和93.67%，F_1 Score和准确率相当；基于EfficientNet B0—B4网络的单一子模型的Top-1准确率最高可达96.65%，F_1 Score为96.71%，而Ensemble-Soft模型、Ensemble-Hard模型以及Ensemble-Weight模型的准确率分别为99.07%、98.91%和99.23%，相较于单一子模型，精度进一步提高，F_1 Score与准确率基本相同，模型性能显著；在公开数据集Oxford Flowers102上进行对比试验，3个集成模型相比5个子模型准确率和F_1 Score最高提升了4.56%和5.05%，最低也提升了1.94%和2.29%，证明了本研究提出的迁移和集成学习策略能够有效提高模型性能。［**结论**］本方法可提高荒漠植物的识别准确率，通过云端传输至服务器后，实现荒漠植物的准确识别，为真实野外环境下植物图像识别精度低、模型鲁棒性及泛化性弱等问题提供解决思路。服务于野外调查、教学科普以及科学实验等场景。

关键词：荒漠植物识别；自然背景；集成学习；迁移学习；投票法；数据集

中图分类号：TP183;J522.3　　**文献标志码：**A　　**文章编号：**SA202305001

引用格式：王亚鹏, 曹姗姗, 李全胜, 孙伟. 融合迁移学习和集成学习的自然背景下荒漠植物识别方法[J]. 智慧农业(中英文), 2023, 5(2): 93-103.

WANG Yapeng, CAO Shanshan, LI Quansheng, SUN Wei. Desert plant recognition method under natural background incorporating transfer learning and ensemble learning[J]. Smart Agriculture, 2023, 5(2): 93-103.

1　引　言

荒漠植被维持着荒漠生态系统的物质与能量循环，在控制风沙、防止进一步荒漠化和改善小气候方面具有重要作用，并为人类提供了药材、饲料、木材等副产品，是荒漠生态系统的核心，具有重要的生态和经济意义[1]。目前，因载畜量过高、过度开发和气候变化等因素，许多荒漠植物尚未得到人类深入认识和保护利用，惨遭破坏或灭绝[2,3]，尤其是一些国家级重点保护野生植物，如沙冬青、盐桦等，处于濒危的状态[4]。各类资源调查是荒漠植物研究、保护和利用的长期性、基础性工作，在野外作业过程中，准确、快速判断植物的学名、科属及性状等信息至关重要，但囿于调查人员的技术水平和主观经验，植物分类判别的准确性和一致性存在困难，而且费时费力。

近年来，计算机硬件技术的快速发展以及如智

收稿日期：2023-05-03

基金项目：国家自然科学基金项目（32271880，31860180，32060321）

作者简介：王亚鹏，研究方向为卷积神经网络和图像分类。E-mail：1033453569@qq.com

*通信作者：孙　伟，博士，副研究员，研究方向为农林时空信息智能分析。E-mail：sunwei02@caas.cn

能手机、数码相机、无人机等地面和近地面多源数据获取手段的丰富，以卷积神经网络[5,6]（Convolutional Neural Networks，CNN）为代表的深度学习算法在植物图像分类和表征识别方面取得了突破性进展[7]。目前，有关植物分类研究多在实验室环境下以植物叶片或局部器官为主要研究对象[8-10]，能够获得较高的识别精度，但难以实际应用到自然环境下整株植物图像识别上，同时例如光照、土壤、阴影以及其他植被造成的错综复杂的生长环境将提高识别难度[11]，模型的泛化能力较差且识别准确率受到影响，如何在真实的环境下保证植物的识别准确率将成为研究热点和难点。冯海林等[12]提出了一种基于树木整体图像的树种识别方法，相比于以植物叶片为研究对象的识别精度有较大提升，识别准确率达到99.15%。宋晓宇等[13]以植物的叶片和花朵为研究对象，提出一种有效区域筛选的植物图像识别方法，有效提高了识别准确率，为解决自然背景下植物图像识别问题提供了新思路。Zhou等[14]依据6种自然复杂背景下蔬菜疾病图像，提出了一种基于区域建议和渐进式学习的蔬菜疾病识别方法，可引导模型聚焦于具有自然复杂背景的疾病图像中的关键区域，避免手动标记图像的昂贵成本，识别准确率达到98.26%。Li等[15]以24种荒漠植物为研究对象，从不同角度比较了一系列性能良好的CNN模型，得出MobileNet V2网络在精度、参数量和浮点运算数（Floating Point Operations，FLOPs）之间实现了最佳平衡。目前，解决自然背景下的植物图像识别的关键在于提高属于同一科、属下的识别准确率[16]，以及克服各种影响精度的天然噪声[17]。

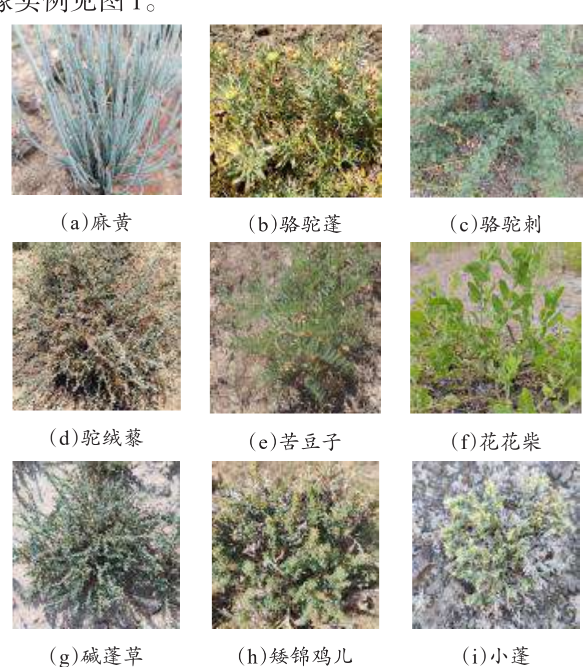
本研究为解决自然环境下整株荒漠植物图像识别问题，提出一种基于迁移学习和集成学习的荒漠植物识别方法。为最大化利用荒漠植物图像，使用K折交叉验证划分数据集，而后将K个子模型在源数据ImageNet上进行迁移学习，使用投票法集成模型，实现对荒漠植物的准确识别。用智能识别算法辅助传统人工识别荒漠植物，能够为荒漠植物认识及多样性保护，固沙、饲料及药用植物开发与应用，优良薪炭材开发等领域提供技术支持，并减少对专家鉴定的依赖，节省时间和人力成本，为野外调查人员提供支持，为保护和利用自然植物资源提供科学依据。

2　材料与方法

2.1 数据采集与预处理

2.1.1　数据集构建

分别集中于2021年9月底和2022年7月至8月，在新疆昌吉及塔城地区采集荒漠植物图像数据。使用智能手机和数码相机拍摄野外荒漠植物图像，像素大小分别为3968×2232和4800×3200。采取多角度多光线拍摄，拍摄时保证拍摄的荒漠植物图像占据图像中央的主要部分，且包含例如阴影、砂石、光照及其他杂类草遮挡等噪声，以增强数据集的鲁棒性。同时记录每种荒漠植物的专家鉴定结果，构建荒漠植物图像数据集，共计12,507张图像，包括13科、43属、50种植物类型，每种荒漠植物图像数量在183～339张，荒漠植物图像实例见图1。



(a)麻黄　(b)骆驼蓬　(c)骆驼刺
(d)驼绒藜　(e)苦豆子　(f)花花柴
(g)碱蓬草　(h)矮锦鸡儿　(i)小蓬

图1　数据集中荒漠植物图像实例

Fig. 1 Examples of desert plant images in data set

2.1.2　数据预处理

由于图像原图较大，需将图像大小统一调整为600×600。以2∶8的比例将数据集划分为测试集（2481张）和训练集（10,026张），测试集在进行测试之前不进行数据增强操作，以保证数据的独立性

和无重复性。而后通过五折交叉验证将训练集划分为4个训练子集和1个验证子集，因训练子集图像数量较少，且难以覆盖真实野外的多种复杂情况，导致模型在训练过程中可能出现过度拟合，在验证集上表现出较高的识别准确率，但在测试集上的识别效果不佳，模型的泛化性和鲁棒性较差。因此对于训练子集，在训练过程中需进行数据增强操作。通过分析荒漠植物表征，采取图像平移、翻转、旋转、随机颜色以及随机黑色块遮挡（黑色块长边区间为［100，600］，短边区间为［10，60］）5种方法进行数据增强，且过程中随机选择、组合，设置不同的触发概率，将训练子集图像数量扩充至5倍。以草麻黄为例，数据增强实例见图2，其中图2（a）为原图，图2（b）～图2（f）为单一方法的效果图，图2（g）～图2（i）实际上为5种数据增强方法随机选择、组合后的效果图。

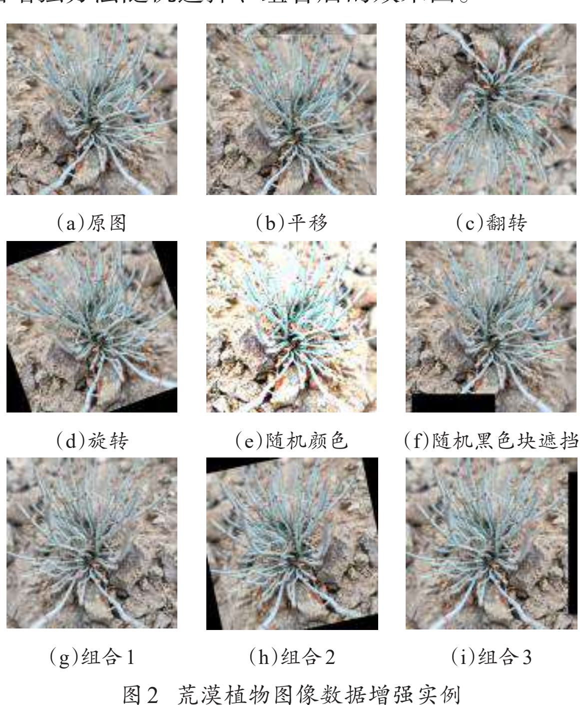


图2 荒漠植物图像数据增强实例

Fig. 2 Data enhancement instance of desert plant images

2.2 EfficientNet网络

根据工程经验，若增大网络输入层的图像分辨率，则需要更深、更宽的卷积层来增加感受野、捕捉图像特征，复合放大神经网络的深度、宽度及分辨率等可以使模型更好地拟合图像的特征，从而获得更好的识别效果[18, 19]。通过搜索网络的深度、宽度以及图像的分辨率，达到参数合理化配置的目的，以此来开发基础网络EfficientNet B0[20]，复合放大三者构建B1—B7网络。

EfficientNet B0共分为9个阶段。第1阶段是通过3×3卷积进行下采样操作，第2至8阶段通过堆叠7个MBConv Block，达到提取图像特征的目的，最后连接1个1×1卷积、Global average pooling（全局平均池化）和1个全连接层，EfficientNet B0网络架构见表1。

表1 EfficientNet B0网络架构

Table 1 EfficientNet B0 network architecture

操作	Input(224×224 RGB Image)				
	卷积核大小	步距	倍率因子N	输出通道	输出尺寸
Conv×1	3×3	2	1	32	112×112
MBConv×1	3×3	1	6	16	112×112
MBConv×2	3×3	2	6	34	56×56
MBConv×2	5×5	2	6	40	28×28
MBConv×3	3×3	2	6	80	14×14
MBConv×3	5×5	1	6	112	14×14
MBConv×4	5×5	2	6	192	7×7
MBConv×1	3×3	1	6	320	7×7
Conv1×1 & Global average pooling & FC					

EfficientNet B0网络的输入图像大小为工程中常用的224×224，通过堆叠MBConv Block来增加网络的深度和宽度，而B1—B7网络是在B0的基础上放大深度、宽度和图像的分辨率。MBConv Block结构见图3。

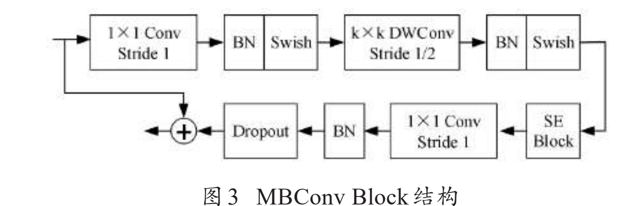


图3 MBConv Block结构

Fig. 3 Structure of MBConv Block

2.3 迁移学习

广义上讲，利用已有的知识、模型、结构达成在目标数据上的学习目标，这一过程称为迁移学习[21]。预训练及微调（Pre-Train and Fine-Tune）长久以来被认为是迁移学习最重要的表现形式，是指在源领域训练好一个网络，直接将其用于目标域的数据，并在目标域上进行微调。本研究基于EfficientNet B0—B4网络进行迁移学习，训练过程中冻

结1至8阶段的权重，只训练最后一个1×1卷积层和全连接层，在图像数量较少的情况下，也可快速训练出理想的模型。

2.4 集成学习

2.4.1 模型架构

在数据充足时，经典的机器学习方法不如深度学习方法有效，但集成学习（Ensemble Learning）在深度学习领域中仍然发挥着巨大的作用[22]。目前，在深度学习方法之上引入集成学习仍然是许多深度学习研究者来提升效果的重要手段。集成学习与其他深度学习方法的核心区别在于其专注于偏差、方差权衡（Bias-Variance Tradeoff）问题，使得集成学习对包括深度学习在内的所有机器学习方法都具有价值。在统计学习中，衡量一个模型的好坏主要通过偏差和方差，一般认为低偏差和低方差是理想的结果，集成学习的偏差及方差示意见图4。

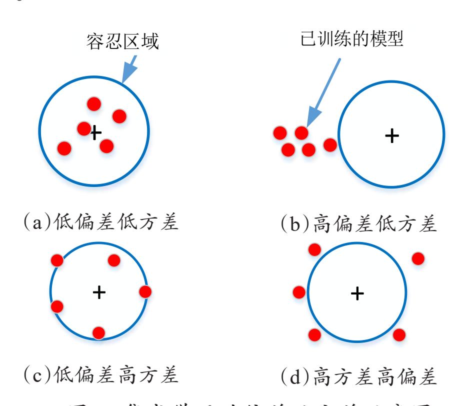


图4 集成学习的偏差及方差示意图

Fig. 4 Diagram of bias and variance of ensemble learning

如图4所示，若模型表现出高偏差、高方差以及高偏差高方差三类情况，则需要通过策略降低偏差和方差。当偏差过高时，可以通过增加模型复杂度、Boosting[23]和Stacking[24]等方法减少偏差。当方差过高时，可以通过降低模型复杂度、Bagging[25]和Stacking等方法减少方差。集成学习的核心思想是通过训练多个模型，并通过一定的方式结合起来，能够降低偏差和方差，从而提升模型性能。

本研究基于EfficientNet网络，首先通过迁移学习在ImageNet数据集上进行预训练，而后采用将Bagging和Stacking结合的集成学习策略，共分为两层。第一层借鉴Stacking方法引入K折交叉验证划分数据集以及训练K个子模型，考虑到本研究中分类问题每种模型的输出特征相同，第二层采用Bagging的方式将第一层模型通过投票法集成，区别是采用一个基学习器和K个基学习器进行对比，选取更优的基学习器，以此构建集成模型，降低模型偏差和方差，提升模型的识别性能。荒漠植物图像识别集成模型架构见图5。

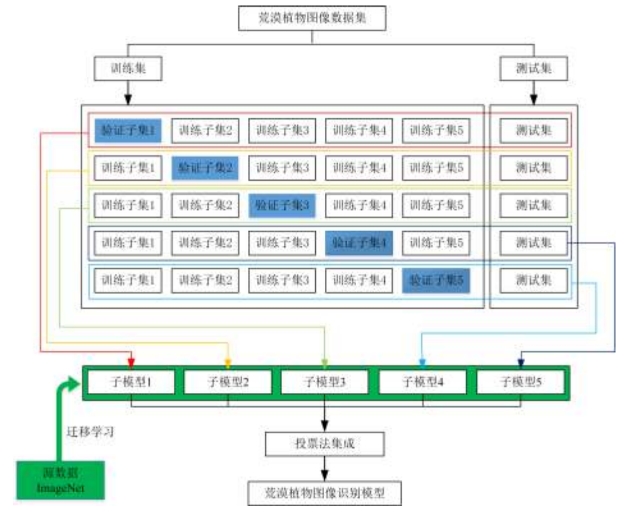


图5 荒漠植物图像识别集成模型架构

Fig. 5 Integrated model architecture of desert plant image recognition

2.4.2 模型集成策略

投票法是一种遵循少数服从多数原则的集成学习策略，通过集成多个模型降低方差，从而提高模型的鲁棒性，达到高准确率的目的。在理想情况下，投票法的预测结果应当优于任何单一模型的预测结果。针对本研究所开展的荒漠植物图像分类问题，投票法预测的结果是所有模型中出现最多的预测结果，分类投票法又分为软投票和硬投票，见表2。

表2 软投票和硬投票的区别

Table 2 Difference of soft voting and hard voting

模型	样本	
	预测类别	预测为A的概率/%
子模型1	A类别	91
子模型2	B类别	49
子模型3	B类别	49
硬投票集成	B类别	
软投票集成	A类别	

如表2所示，对于同一个样本，软投票（Ensemble-Soft）与硬投票（Ensemble-Hard）可以得出不同的结论。对于一个简单的二分类，一个样本的结果是A或者B，三个子模型对于这个样本分别有三个识别结果（以及预测概率）。对于硬投票而言，遵循少数服从多数原则，而软投票则是考虑预测概率这一额外信息，将概率进行平均，在这个例子中，超过50%则认为是A类，否则是B类。相对于硬投票，软投票考虑到预测概率这一额外信息，因此可以得出比硬投票法更加准确的预测结果。考虑到集成模型过程中各个子模型对荒漠植物图像表征提取能力的不同，出现识别准确率差异的情况，应采取更为合理的加权投票法（Ensemble-Weight）集成模型。第 i 个子模型的预测分类结果如矩阵 $\boldsymbol{J}$ 公式（1）所示。

$$\boldsymbol{J}=\begin{bmatrix} P(y_i=1) \\ P(y_i=2) \\ \cdots \\ P(y_i=n) \end{bmatrix} \tag{1}$$

其中，$i=1，2，\cdots，5$；$n=1，2，\cdots，50$。该矩阵每一行代表子模型对荒漠植物图像的预测概率，模型集成后可通过加权取各个子模型的平均值得到预测结果 $\boldsymbol{E}$，如公式（2）所示。

$$\boldsymbol{E}=\sum_{i=1}^{5} w_i \boldsymbol{J} \tag{2}$$

其中，$w_i=\dfrac{m_i}{1+2+3+4+5}$；$m_i=1，2，\cdots，5$；$m_i$ 的取值依据在测试集上各个子模型识别准确率的高低顺序，准确率最低的子模型 $m_i=1$，准确率最高的子模型 $m_i=5$，以此通过加权投票法得到预测结果 $\boldsymbol{E}$。

2.5 模型评价指标

2.5.1 Top-1 准确率

图像分类问题使用测试集 $top(k)$ 准确率评价模型的优劣，其计算如公式（3）。

$$top(k)=\frac{r}{a} \tag{3}$$

其中，$k=1，2，\cdots，5$；a 表示模型测试集图像的总数量，r 表示模型预测的前 k 个结果中测试正确的图像数量。一般常用Top-1准确率，是指排名第一的类别与实际结果相符的准确率。

2.5.2 分类指标

混淆矩阵是评判模型结果的一种指标，常用于评判分类器模型的优劣[26]。在混淆矩阵的基础上延伸了精确率（*Precision*）、召回率（*Recall*）以及 F_1 Score等指标，预测性分类模型，期望越准确越理想，对应到混淆矩阵中，期望TP（True Positive）与TN（True Negative）的数量大，而FP（False Negative）与FN（False Positive）的数量小。F_1 Score的物理意义是将精确率和召回率进行加权平均，公式（6）认为两者权重相等，其取值范围从0到1，1代表模型的输出结果最好，0代表模型的输出结果最差。混淆矩阵的分类指标如表3所示。

表3 混淆矩阵的分类指标

Table 3 Classified indicators of Confusion matrix

指标	意义
$Precision=\dfrac{TP}{TP+FP}$ （4）	模型预测为正类的样本中，预测正确的比例
$Recall=\dfrac{TP}{TP+FN}$ （5）	模型正确预测为正类的样本数占总的正类样本数的比例
$F_1\,\text{Score}=\dfrac{2Precision\times Recall}{Precision+Recall}$ （6）	综合了Precision与Recall产出的结果，认为两者同样重要

2.6 实验环境及参数设置

为确保模型运行环境的一致性，消除实验结果因环境因素带来的影响，本实验所有模型训练及测试过程中均采用相同的实验环境，具体如表4所示。

表4 模型训练及测试的实验环境

Table 4 Experimental environment for model training and testing

软件	硬件
编译器：Pycharm 2021.1.1	处理器：Intel i7-10750H CPU
语言：Python 3.7.0	内存：16 G RAM
深度学习框架：PyTorch 1.11.0	图形处理器：NVIDIA GeForce GTX 1660 Ti
运算平台：CUDA 10.0.130	显存：6 G

模型训练中，每次训练所选取的样本数Batch Size设为32，迭代次数Epoch设为100，损失函数为交叉熵损失函数，优化器选用随机梯度下降，采用余弦退火策略设置学习率，初始学习率设为0.01，100个Epoch后减小到0.0001。

3 结果与分析

3.1 基于EfficientNet B0网络的荒漠植物识别结果

研究对象为50种荒漠植物，以EfficientNet B0为基础网络，基于迁移学习和集成学习，构建荒漠植物图像识别模型。根据50类荒漠植物图像数量，首先划分20%的数据作为测试集，剩余采用五折交叉验证划分数据集，则可用DPi (i = 1，2，···，5)表示每一折训练集和验证集。在测试集上，基于EfficientNet B0网络的荒漠植物识别结果如表5所示。

表5 EfficientNet B0网络的荒漠植物识别结果

Table 5 Desert plant recognition results using EfficientNet B0 network

模型	Top-1准确率/%	精确率/%	召回率/%	F_1 Score/%
EfficientNet B0-DP1	92.99	93.66	92.89	93.28
EfficientNet B0-DP2	92.62	93.48	92.54	93.01
EfficientNet B0-DP3	92.26	92.95	92.21	92.58
EfficientNet B0-DP4	93.23	93.61	93.17	93.39
EfficientNet B0-DP5	93.35	93.85	93.27	93.56
Ensemble-Soft	93.63	94.24	93.52	93.88
Ensemble-Hard	93.55	94.12	93.44	93.78
Ensemble-Weight	93.67	94.25	93.57	93.91

如表5所示，经过5个不同训练集训练出的EfficientNet B0模型，对于同一份测试集5个子模型的Top-1准确率各不相同，最高为93.35%，最低为92.26%，相差1.09%，表明经过五折交叉验证划分后，对于同一网络表现出不同的识别准确率，数据集存在一定的差异性。5个EfficientNet B0模型的F_1 Score也均有差异，最高为93.56%，最低为93.01%，相差0.55%，Top-1准确率与F_1 Score的值相当，可见模型识别准确率达到较高水平，性能较为显著。从精确率和召回率的值可以看出，对于5个子模型和3个集成模型，精准率都高于Top-1准确率，表明模型对正样本结果中的预测准确程度高于整体的预测准确程度，同理，召回率都低于Top-1准确率，表明模型对实际为正的样本预测为正样本的比例低于整体的预测准确程度，但两者相差不大，都达到了较高的水平，而F_1 Score综合了两类结果，更为全面地评价了模型的性能优劣。

通过集成学习和投票法策略集成模型后，模型的Top-1准确率和F_1 Score都有了提升，Ensemble-Soft模型准确率相对于EfficientNet B0-DP3模型最高提升了1.37%，F_1 Score提升1.3%，而软投票相比硬投票策略，准确率和F_1 Score略高，验证了软投票考虑预测概率，可以得出更加准确的预测结果。通过加权投票法集成的Ensemble-Weight模型，相比于Ensemble-Soft模型和Ensemble-Hard模型的准确率和F_1 Score提升不明显，但表明本研究提出的加权投票法效果优于二者。3个集成模型对于EfficientNet B0-DP5模型，准确率和F_1 Score提升不大，可得出结论，对于5个EfficientNet B0模型，虽模型的识别效果相差不大，但模型之间较高的同质性限制了投票法策略的发挥，甚至对于集成模型而言，EfficientNet B0-DP5模型是其识别效果是否显著的关键所在，所以应考虑使用差异性较大的网络作为子网络。5个EfficientNet B0网络在验证集上的损失值变化曲线如图6所示。

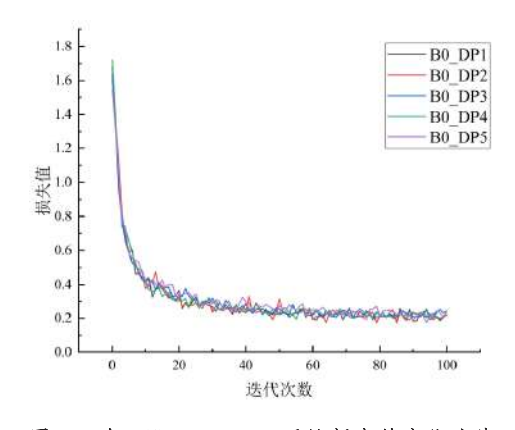


图6 5个EfficientNet B0网络损失值变化曲线

Fig. 6 Loss value variation curves of five EfficientNet B0 networks

3.2 基于差异化网络的荒漠植物识别结果

EfficientNet B1—B7网络是在EfficientNet B0的基础上对网络的深度、宽度和图像的分辨率进行放大操作，且在MBConv Block中捷径连接和主分支相加之前，进行Dropout操作，在防止网络过度拟合的过程中会随机丢弃整个MBConv Block，只剩下捷径连接，这使得模型在训练过程中具有随机

深度，进一步放大了EfficientNet系列网络间的差异性，使得EfficientNet系列网络可作为子模型用于集成学习。

在相同的训练条件下，考虑到模型的识别效率问题选用EfficientNet B0—B4网络作为子模型，B0网络保持不变，将 DPi (i = 2，3，4，5) 每一折数据集作为EfficientNet B1—B4网络的训练集和验证集，基于迁移学习和集成学习，构建荒漠植物图像识别模型，基于差异化网络的荒漠植物识别结果见表6。

表6 差异化网络的荒漠植物识别结果

Table 6 Recognition results of desert plants based on differential networks

模型	Top-1准确率/%	精确率/%	召回率/%	F_1 Score/%
EfficientNet B0-DP1	92.99	93.66	92.89	93.28
EfficientNet B1-DP2	93.43	93.84	93.27	93.55
EfficientNet B2-DP3	95.45	95.63	95.45	95.54
EfficientNet B3-DP4	96.57	96.71	96.53	96.62
EfficientNet B4-DP5	96.65	96.77	96.66	96.71
Ensemble-Soft	99.07	99.06	99.07	99.07
Ensemble-Hard	98.91	98.93	98.90	98.91
Ensemble-Weight	99.23	99.24	99.23	99.23

如表6所示，对于同一份测试集Top-1准确率最高为96.65%，F_1 Score为96.71%，两者相当，可见EfficientNet B4单一网络已能达到较高的精度，模型性能显著。因EfficientNet B1—B4网络基于B0网络在深度、宽度和分辨率三个尺度上都进行了放大操作，Top-1准确率和F_1 Score皆有提升，证明了随着三者的不断扩大，可达到更高的模型识别准确率。但同时注意到，F_1 Score相比准确率提升放缓，且EfficientNet B4-DP5模型的精度已经略高于EfficientNet B3-DP4，说明此时继续放大三者可能会出现模型退化现象，在训练难度加大的情况下，若出现模型退化现象会使成本增大，更加不利于集成模型。通过对比精确率和召回率，呈现与集成5个EfficientNet B0网络大致相同的结果，精确率高于Top-1准确率，而召回率会出现与Top-1准确率相当的情况，此时模型更加注重对实际为正的样本预测为正样本的结果，说明在一定程度上放大三者能够得到更高的精确率和召回率，F_1 Score也随之提高。

经过投票法策略集成模型后，Ensemble-Soft模型Top-1准确率和F_1 Score相比于EfficientNet B0-DP1提升了6.08%和5.79%，提升效果显著，相比于EfficientNet B4-DP5，Ensemble-Soft模型也有2.42%和2.36%的提高，对比集成5个EfficientNet B0网络性能提升效果显著，说明投票法策略更适用于集成学习中模型同质性低的子网络。软投票相比硬投票策略，准确率和F_1 Score都提升了0.16%，再次验证了软投票效果更佳。通过加权投票法集成的Ensemble-Weight模型，准确率和F_1 Score软投票和硬投票也有了提升，均达到99.23%，模型性能显著且识别效果更佳，再次表明加权投票法效果优于二者。

另外，基于Ensemble-Weight模型，测试采集到的50种植物的Top-1准确率，均达到了97%及以上。EfficientNet B0—B4网络在验证集上的损失值变化曲线如图7所示。

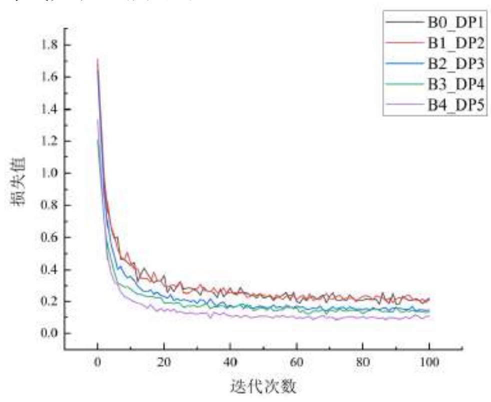


图7 EfficientNet B0—B4网络损失值变化曲线

Fig. 7 Loss value variation curves of EfficientNet B0–B4 network

3.3 集成学习的有效性验证

为验证本研究所提出的集成学习方法的有效性，采用公开数据集进行对比。Oxford Flowers102数据集是英国常见的102种花卉植物的一类数据，每种花卉植物包含40至258张图像，共计8189张图像，该花卉植物图像数据集在姿势和光线上具有较大比例的变化。此外，部分类别在该类别和几个非常相似的类别中有很大的差异。考虑到Oxford Flowers102数据集的公开性，与本研究所建立的荒漠植物图像数据集在形态特征分布上具有一定的相

似性，种类较多且在数量分布上具有较大的差异性，可以作为本研究提出的集成学习方法有效性验证的对比数据集，Oxford Flowers102图像实例见图8。



（a）barbeton daisy （b）sword lily （c）wild pansy （d）orange dahlia （e）gazania （f）water lily

图8 Oxford Flowers102图像实例

Fig. 8 Examples of Oxford Flowers102 images

根据102类花卉植物图像数量，首先划分出20%的测试集，而后采用五折交叉验证划分剩余数据集，则可用OFi（i = 1，2，···，5）表示每一折训练集和验证集。模型训练中，采取与荒漠植物图像识别模型相同的训练策略。在测试集上，基于差异化网络的Oxford Flowers102识别结果见表7。

表7 差异化网络的Oxford Flowers102识别结果

Table 7 Recognition results of oxford flowers102 based on differential network

模型	Top-1准确率/%	精确率/%	召回率/%	F_1 Score/%
EfficientNet B0-OF1	93.13	93.25	91.96	92.60
EfficientNet B1-OF2	93.51	94.32	92.67	93.48
EfficientNet B2-OF3	94.19	94.38	93.34	93.85
EfficientNet B3-OF4	94.76	94.57	94.60	94.58
EfficientNet B4-OF5	95.13	95.15	94.08	94.61
Ensemble-Soft	97.63	97.72	97.28	97.50
Ensemble-Hard	97.07	97.06	96.74	96.90
Ensemble-Weight	97.69	97.81	97.49	97.65

如表7所示，5个子模型在Oxford Flowers102和荒漠植物数据集的识别效果类似，Top-1准确率和F_1 Score相当，达到了较高的水平。而3个集成模型相比5个子模型准确率和F_1 Score最高提升了4.56%和5.05%，最低也提升了1.94%和2.29%，验证了本研究提出的集成学习策略能够提高模型的识别准确率，提升模型性能。可以观察到，5个子模型的F_1 Score提升放缓，若继续增大模型体积可能会出现退化现象，训练成本提高且降低识别效率，不利于模型集成。

4 讨论与结论

4.1 讨论

目前，有关于植物图像识别大多局限于单一背景下或实验室环境下的植物识别研究，且识别对象多为植物的叶片或局部器官，但在实际应用中搜集到的大多数植物图像都是自然背景下的，包含例如光照、土壤、杂草、其他植物以及植物成群生长等复杂噪声，因此自然背景下的整株荒漠植物无法运用传统的方式去识别。本研究致力于解决自然背景下的荒漠植物图像识别以及克服各种影响精度的背景干扰和天然噪声，提高精度的同时增加模型的泛化能力。

现阶段，国内外对于荒漠植物的分类识别，主要依靠传统人工识别及专家经验，耗时且费力。用智能识别算法辅助识别荒漠植物，能够有效节省时间和人力成本，并为野外调查人员提供支持和减少对专家鉴定的依赖。本研究构建了自然背景下新疆昌吉及塔城地区50种荒漠植物图像数据集，每种荒漠植物图像数量在183至339张，共计12,507张图像。文献[27-29]基于植物整体外观特征，但研究对象多为草地植物、田间杂草以及室内盆栽。目前，国内暂无有关荒漠植物图像数据集的构建与识别研究。国外已知有关自然背景下荒漠植物的识别研究工作也相对较少，文献[15]构建了一个包含24种荒漠植物，共计2331张图像的荒漠植物图像数据集，在对比一系列模型后实现了在精度与体积之间的最佳平衡，本研究所构建的数据集与其相比，在植物种类及数量上超过前者，识别准确率也远超前者，并且本研究所使用的荒漠植物图像全部拍摄于野外调查，未与前者一样结合公开数据集或网络爬取，能够较为真实地反映荒漠植物在一段时间内的生长特征。本研究所使用的迁移、集成学习策略，能够有效提高模型的识别准确率，在50种荒漠植物图像数据集的最高准确率已达到99.23%，性能显著，且每种植物分类的误差均不大。但在实验过程中发

现以下问题有待完善和解决：

（1）荒漠植物图像数据集还需进一步补充和完善，可从荒漠植物种类及采集阶段两个方向考虑；

（2）考虑其他集成学习策略或选择轻量化网络作为子网络，来改善难以训练和推理速度慢的问题。

4.2 结论

基于PyTorch深度学习框架，以自然环境及自然背景下新疆昌吉和塔城地区采集的荒漠植物图像为研究对象，提出一种融合迁移学习和集成学习的整株荒漠植物图像识别算法，解决了荒漠植物图像在真实野外环境下识别精度低、模型鲁棒能力及泛化能力弱的问题，可应用于野外实地调查、教学科普以及科学实验等场景，弥补了当前国内关于荒漠植物图像数据集构建及机器视觉分类识别研究的不足。本研究基于EfficientNet系列网络，通过迁移学习在ImageNet数据集上进行预训练，而后采用将Bagging和Stacking结合的集成学习策略，第一层采用Stacking的方法并引入K折交叉验证划分数据集以及训练模型，第二层采用Bagging的方式将第一层模型的输出特征通过投票法集成模型，从而得到完整的集成学习模型。

基于EfficientNet B0网络的单一子模型的Top-1准确率最高可达93.35%，最低为92.26%，单一子模型已能够达到较高的精度。通过投票法集成模型后，软投票、硬投票以及加权投票法的准确率分别为93.63%、93.55%和93.67%，F_1 Score和准确率相当，比起单一子模型提高不明显，子模型之间较高的同质性限制了投票法策略的发挥。在使用差异较大的网络作为子模型时，即EfficientNet B0—B4网络，基于EfficientNet B0—B4网络的单一子模型的Top-1准确率最高可达96.65%，F_1 Score为96.71%，模型性能优异。而软投票、硬投票以及加权投票法的准确率分别为99.07%、98.91%和99.23%，相较于单一子模型，精度进一步提高，F_1 Score与准确率基本相同，模型性能显著。在公开数据集Oxford Flowers102上进行了集成学习的有效性验证，发现本研究提出的集成学习策略相比单一子模型能够有效提高模型的识别准确率，提升模型性能。本研究使用深度学习算法识别荒漠植物，能够为荒漠植物认识及多样性保护和野外调查人员提供技术支持，为保护和利用自然植物资源提供科学依据。

利益冲突声明： 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 宋智芳. 伊犁绢蒿荒漠草地植被特征对放牧干扰的响应[D]. 乌鲁木齐: 新疆农业大学, 2018.
SONG Z F. Response of *Seriphidium transiliense* vegetation characteristics to grazing disturance in desert grasslands[D]. Urumqi: Xinjiang Agricultural University, 2018.

[2] 滕迎凤. 宁夏沙湖自然保护区植物多样性研究[D]. 银川: 宁夏大学, 2013.
TENG Y F. Studies on diversity of the plants in Shahu nature reserve, Ningxia, China[D]. Yinchuan: Ningxia University, 2013.

[3] 燕辉. 西北旱区两种典型沙生植物对盐胁迫响应的研究[D]. 杨凌: 西北农林科技大学, 2012.
YAN H. The response of two representative desert shrubs to salt stress in northwest arid region[D]. Yangling: Northwest A & F University, 2012.

[4] 何恒斌. 沙冬青群落及其根瘤菌的研究[D]. 北京: 北京林业大学, 2008.
HE H B. Studies on communities and rhizoibum of *Ammopiptanthus mongolicus* (maxim.)[D]. Beijing: Beijing Forestry University, 2008.

[5] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[6] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning[M]. Cambridge, Massachusetts: The MIT Press, 2016.

[7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[8] JEON W S, RHEE S Y. Plant leaf recognition using a convolution neural network[J]. The international journal of fuzzy logic and intelligent systems, 2017, 17(1): 26-34.

[9] LEE S H, CHAN C S, WILKIN P, et al. Deep-plant: Plant identification with convolutional neural networks[C]// 2015 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2015: 452-456.

[10] 韩斌, 曾松伟. 基于多特征融合和卷积神经网络的植物叶片识别[J]. 计算机科学, 2021, 48(S1): 113-117.
HAN B, ZENG S W. Plant leaf image recognition based on multi-feature integration and convolutional neural network[J]. Computer science, 2021, 48(S1): 113-117.

[11] 金莉婷. 基于卷积神经网络的复杂背景植物图像识别研究[D]. 兰州: 兰州交通大学, 2020.
JIN L T. Research on plant image recognition with complex background based on convolution neural network[D]. Lanzhou: Lanzhou Jiaotong University, 2020.

[12] 冯海林, 胡明越, 杨垠晖, 等. 基于树木整体图像和集成

迁移学习的树种识别[J]. 农业机械学报, 2019, 50(8): 235-242, 279.
FENG H L, HU M Y, YANG Y H, et al. Tree species recognition based on overall tree image and ensemble of transfer learning[J]. Transactions of the Chinese society for agricultural machinery, 2019, 50(8): 235-242, 279.
[13] 宋晓宇, 金莉婷, 赵阳, 等. 基于有效区域筛选的复杂背景植物图像识别方法[J]. 激光与光电子学进展, 2020, 57(4): 181-191.
SONG X Y, JIN L T, ZHAO Y, et al. Plant image recognition with complex background based on effective region screening[J]. Laser & optoelectronics progress, 2020, 57(4): 181-191.
[14] ZHOU J, LI J X, WANG C S, et al. A vegetable disease recognition model for complex background based on region proposal and progressive learning[J]. Computers and electronics in agriculture, 2021, 184: ID 106101.
[15] LI J C, SUN S D, JIANG H R, et al. Image recognition and empirical application of desert plant species based on convolutional neural network[J]. Journal of arid land, 2022, 14(12): 1440-1455.
[16] 曹香滢, 孙卫民, 朱悠翔, 等. 基于科优先策略的植物图像识别[J]. 计算机应用, 2018, 38(11): 3241-3245.
CAO X Y, SUN W M, ZHU Y X, et al. Plant image recoginiton based on family priority strategy[J]. Journal of computer applications, 2018, 38(11): 3241-3245.
[17] 郭晓丽. 基于全卷积神经网络的植物图像分割算法研究与实现[D]. 呼和浩特: 内蒙古大学, 2021.
GUO X L. Research and implementation on plant image segementation algorithm based on neural network[D]. Hohhot: Inner Mongolia University, 2021.
[18] RAGHU M, POOLE B, KLEINBERG J, et al. On the expressive power of deep neural networks[C]// Proceedings of the 34th International Conference on Machine Learning -Volume 70. New York, USA: ACM, 2017: 2847-2854.
[19] ZAGORUYKO S, KOMODAKIS N. Wide residual networks[EB/OL]. arXiv: 1605.07146, 2016.
[20] TAN M, LE Q. EfficientNet: Rethinking model scaling for convolutional neural networks[EB/OL]. International conference on machine learning. arXiv:1905.11946, 2019.
[21] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE transactions on knowledge and data engineering, 2010, 22(10): 1345-1359.
[22] DONG X B, YU Z W, CAO W M, et al. A survey on ensemble learning[J]. Frontiers of computer science, 2020, 14(2): 241-258.
[23] WANG B, PINEAU J. Online bagging and boosting for imbalanced data streams[J]. IEEE transactions on knowledge and data engineering, 2016, 28(12): 3353-3366.
[24] HUI Y, MEI X S, JIANG G D, et al. Milling tool wear state recognition by vibration signal using a stacked generalization ensemble model[J]. Shock and vibration, 2019, 2019: 1-16.
[25] ANDIOJAYA A, DEMIRHAN H. A bagging algorithm for the imputation of missing values in time series[J]. Expert systems with applications, 2019, 129: 10-26.
[26] FIELDING A H, BELL J F. A review of methods for the assessment of prediction errors in conservation presence/absence models[J]. Environmental conservation, 1997, 24(1): 38-49.
[27] 高宏元, 高新华, 冯琦胜, 等. 基于深度学习的天然草地植物物种识别方法[J]. 草业科学, 2020, 37(9): 1931-1939.
GAO H Y, GAO X H, FENG Q S, et al. Approach to plant species identification in natural grasslands based on deep learning[J]. Pratacultural science, 2020, 37(9): 1931-1939.
[28] 彭文, 兰玉彬, 岳学军, 等. 基于深度卷积神经网络的水稻田杂草识别研究[J]. 华南农业大学学报, 2020, 41(6): 75-81.
PENG W, LAN Y B, YUE X J, et al. Research on paddy weed recognition based on deep convolutional neural network[J]. Journal of South China agricultural university, 2020, 41(6): 75-81.
[29] 陈淑君, 周永霞, 方勇军. 基于整体外观特征的植物种类识别研究[J]. 计算机应用与软件, 2017, 34(9): 222-227.
CHEN S J, ZHOU Y X, FANG Y J. The plant species recognition based on the whole appearanc features[J]. Computer applications and software, 2017, 34(9): 222-227.

Desert Plant Recognition Method Under Natural Background Incorporating Transfer Learning and Ensemble Learning

WANG Yapeng[1,2], CAO Shanshan[2,3], LI Quansheng[1], SUN Wei[2,3*]

(1. *Computer and Information Engineering College*, *Xinjiang Agricultural University*, *Urumqi* 830052, *China*; 2. *Agricultural Information Institute*, *Chinese Academy of Agricultural Sciences*, *Beijing* 100081, *China*; 3. *National Agriculture Science Data Center*, *Beijing* 100081, *China*)

Abstract:

[Objective] Desert vegetation is an indispensable part of desert ecosystems, and its conservation and restoration are crucial. Accurate identification of desert plants is an indispensable task, and is the basis of desert ecological research and conservation. The complex growth environment caused by light, soil, shadow and other vegetation increases the recognition difficulty, and the generalization abili-

ty is poor and the recognition accuracy is not guaranteed. The rapid development of modern technology provides new opportunities for plant identification and classification. By using intelligent identification algorithms, field investigators can be effectively assisted in desert plant identification and classification, thus improve efficiency and accuracy, while reduce the associated human and material costs.

[Methods] In this research, the following works were carried out for the recognition of desert plant: Firstly, a training dataset of deep learning model of desert plant images in the arid and semi-arid region of Xinjiang was constructed to provide data resources and basic support for the classification and recognition of desert plant images.The desert plant image data was collected in Changji and Tacheng region from the end of September 2021 and July to August 2022, and named DPlants50. The dataset contains 50 plant species in 13 families and 43 genera with a total of 12,507 images, and the number of images for each plant ranges from 183 to 339. Secondly, a migration integration learning-based algorithm for desert plant image recognition was proposed, which could effectively improve the recognition accuracy. Taking the EfficientNet B0—B4 network as the base network, the ImageNet dataset was pre-trained by migration learning, and then an integrated learning strategy was adopted combining Bagging and Stacking, which was divided into two layers. The first layer introduced K-fold cross-validation to divide the dataset and trained K sub-models by borrowing the Stacking method. Considering that the output features of each model were the same in this study, the second layer used Bagging to integrate the output features of the first layer model by voting method, and the difference was that the same sub-models and K sub-models were compared to select the better model, so as to build the integrated model, reduce the model bias and variance, and improve the recognition performance of the model. For 50 types of desert plants, 20% of the data was divided as the test set, and the remaining 5 fold cross validation was used to divide the dataset, then can use $DPi(i=1,2,\cdots,5)$ represents each training or validation set. Based on the pre trained EfficientNet B0—B4 network, training and validation were conducted on 5 data subsets. Finally, the model was integrated using soft voting, hard voting, and weighted voting methods, and tested on the test set.

[Results and Discussions] The results showed that the Top-1 accuracy of the single sub-model based on EfficientNet B0 network was 92.26%~93.35%, the accuracy of the Ensemble-Soft model with soft voting, the Ensemble-Hard model with hard voting and the Ensemble-Weight model integrated by weighted voting method were 93.63%, 93.55% and 93.67%, F_1 Score and accuracy were comparable, the accuracy and F_1 Score of Ensemble-Weight model integrated by weighted voting method were not significantly improved compared with Ensemble-Soft model and Ensemble-hard model, but it showed that the effect of weighted voting method proposed in this study was better than both of them. The three integrated models demonstrate no noteworthy enhancements in accuracy and F_1 Score when juxtaposed with the five sub-models. This observation results suggests that the homogeneity among the models constrains the effectiveness of the voting method strategy. Moreover, the recognition effects heavily hinges on the performance of the EfficientNet B0-DP5 model. Therefore, the inclusion of networks with more pronounced differences was considered as sub-models. A single sub-model based on EfficientNet B0—B4 network had the highest Top-1 accuracy of 96.65% and F_1 Score of 96.71%, while Ensemble-Soft model, Ensemble-Hard model and Ensemble-Weight model got the accuracy of 99.07%, 98.91% and 99.23%, which further improved the accuracy compared to the single sub-model, and the F_1 Score was basically the same as the accuracy rate, and the model performance was significant. The model integrated by the weighted voting method also improved accuracy and F_1 Score for both soft and hard voting, with significant model performance and better recognition, again indicating that the weighted voting method was more effective than the other two. Validated on the publicly available dataset Oxford Flowers102, the three integrated models improved the accuracy and F_1 Score of the three sub-models compared to the five sub-models by a maximum of 4.56% and 5.05%, and a minimum of 1.94% and 2.29%, which proved that the migration and integration learning strategy proposed in this paper could effectively improve the model performances.

[Conclusions] In this study, a method to recognize desert plant images in natural context by integrating migration learning and integration learning was proposed, which could improve the recognition accuracy of desert plants up to 99.23% and provide a solution to the problems of low accuracy, model robustness and weak generalization of plant images in real field environment. After transferring to the server through the cloud, it can realize the accurate recognition of desert plants and serve the scenes of field investigation, teaching science and scientific experiment.

Key words: desert plant image classification; natural background; ensemble learning; transfer learning; voting method; dataset